Digital Archiving: A Brief Introduction

Katie L. Coldiron

Florida International University, kcoldiro@fiu.edu

Follow this and additional works at: https://digitalcommons.fiu.edu/wphl-da

Recommended Citation

This work is brought to you for free and open access by FIU Digital Commons. It has been accepted for inclusion in Wolfsonian Public Humanities Lab (WPHL) Digital Archive by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.
DIGITAL ARCHIVING: A BRIEF INTRODUCTION

Spring 2022 Community Data Curation Intern Orientation

Katie Coldiron, Digital Archivist

Wolfsonian Public Humanities Lab-FIU
The Community Data Curation Mellon Grant represents a partnership between the Wolfsonian Public Humanities Lab-FIU and eight distinct organizations around South Florida.

- **Sant La Haitian Neighborhood Center**: Non-profit and community-based service agency with a small archive.
- **Stonewall National Museum & Archives**: Major U.S. LGBTQ archive and library with small exhibition space.
- **World AIDS Museum and Educational Center**: Small, relatively new museum focused on preserving AIDS-related narratives and eliminating stigma associated with HIV/AIDS.
- **African American Research Library and Cultural Center**: Flagship branch of the Broward County Library System focused on the preservation of African, African American, and Caribbean cultures. Includes an archive, library, and exhibition space.
- **Historic Hampton House Museum and Cultural Center**: Former Green Book hotel turned museum, educational and event space. Has a small archive of primarily objects and photos.
- **Museum of Graffiti**: First museum in the world dedicated to graffiti art.
- **Vizcaya Museum and Gardens**: House museum with archives related to construction and maintenance of the estate.
- **Jewish Museum of Florida-FIU**: University-affiliated museum dedicated to documenting the Jewish diaspora of the State of Florida. Archive consists of mainly photos.
What will you be doing?

- No two interns will have the exact same work.
- For all interns, the archival work will include the following, with some nuances:
  - Appraising and preparing materials for digitization
  - Digitizing materials
  - Aiding in ingest of materials into dPanther and/or institution’s digital repository
    - Metadata
    - Photo editing
    - Optical Character Recognition (OCR)
  - Oral history collection
Appraising and Preparing Materials for Digitization

- Appraisal: The process of determining whether records and other materials have permanent (archival) value. - Society of American Archivists Dictionary (2022)

- In conjunction with your on-site supervisor, you will assess archival collections value for digitization.

- Questions to ask when appraising for digitization:
  - Is the item part of a larger series or collection already flagged for digitization?
  - Is the item in a poor physical state?
  - Is the item legacy media? (CD's, videocassettes, cassette tapes, etc.)
  - What is the item’s copyright status?

- Once items are identified, they should be put in an easy-to-follow spreadsheet list (more on that later).
<table>
<thead>
<tr>
<th></th>
<th>Collection 1</th>
<th>Collection 2</th>
<th>Collection 3</th>
<th>Collection 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usage (physical materials) 1=Low to 5=High</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anticipated usage (digital materials) 1=Low to 5=High</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unique content 1=Common to 5=Rare</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copyrights cleared 1=no, 3=in process, 5=cleared</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition 1=excellent to 5=poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within scope mission/collection development 1=no to 5=yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: (Rogers and Urbec 2021)
Copyright: Determining an Item’s Status

• What is copyright?
  • “A type of intellectual property that protects original works of authorship as soon as an author fixes a work in a tangible form of expression.” (U.S. Copyright Office 2022)

• To comply with the legal restrictions of copyright, anyone wishing to use the material must comply with fair use practice as outlined in a Rights Statement or Creative Commons license.

• Four Factors of Fair Use
  • The Purpose and Character of the Use
  • The Nature of the Copyrighted Work
  • The Amount or Sustainability of the Portion Used
  • The Effect of the Use on the Potential Market for or Value of the Work.

• All items that go into dPanther must have Rights Statements.

Sources: (Rogers and Urbec 2021) (Columbia University 2022)
Test Your Copyright Knowledge!

The narrator holds copyright, but release forms allow for publishing and use by other persons and entities.

Materials made and released by U.S. government entities are normally public domain.

Those involved in the creation of the album hold copyright (Lou Rawles et al.)
Digitizing Materials

- Different items require different techniques
  - Paper: Flatbed scanner, baseboard with camera on copy stand, book cradle, foam padding, opaque paper, paperweights, staple remover, etc.
  - Objects: Camera with tripod or copy stand, neutral background, paperweights and other objects to position object, etc.
  - Legacy audiovisual media: Cassette tape player, software like Adobe Audition, VHS player, software like Elgato Video Capture

- Lighting is everything
  - You won’t always be able to control where you are digitizing.
  - Full spectrum lighting is the goal.
  - External lights are checked every day before beginning work with a color target.
  - Not applicable to the flatbed scanner or legacy audiovisual media.

Source: Les Goldschmidt
More Equipment You Might See on the Job

Flatbed Scanner

VHS to Digital Conversion Station

Audio Cassette Tape to Digital Conversion Station
Ingesting Materials into a Digital Repository

• Items are digitally archived once they are in a digital repository.
• Items can be discoverable or restricted, depending on wishes of the creator and/or donor.
• Keys for discoverability
  • Descriptive metadata in a recognized schema/structure
  • Clear photos with alt text and organized PDF files
  • Optical character recognition for items including text

Image Source: dpanther.fiu.edu
What is Metadata?

• “Data about Data”
• 4 Key Standards
  • Structure: Also called schema, different ones are employed for different institutions and disciplines (MARC: LOC standard, Dublin Core, MODS: MARC/Dublin Core hybrid, VRA Core: images and works of art, etc.)
  • Content: Formatting rules (ex. last name, first name)
  • Value: Controlled vocabularies (ex. LOC subject headings, inclusive linked data initiatives like homosaurus.org, Change the Subject movement)
  • Format: How to encode the metadata for machine readability (ex. XML vs. CSV format)

• Metadata in Action
  • FIU Groundbreaking

Information Source: (Cofield 2022)
<table>
<thead>
<tr>
<th>identifier</th>
<th>dateCreate</th>
<th>name</th>
<th>roleTerm3</th>
<th>creator2</th>
<th>roleTerm2</th>
<th>type</th>
<th>creator1</th>
<th>language</th>
<th>roleTerm1</th>
<th>type</th>
<th>language</th>
<th>roleTerm1</th>
<th>creator2</th>
<th>roleTerm2</th>
<th>name</th>
<th>type</th>
<th>roleTerm3</th>
<th>dateCreate</th>
</tr>
</thead>
<tbody>
<tr>
<td>hal000001</td>
<td>2017-07-17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hal000002</td>
<td>2017-07-17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hal000003</td>
<td>2017-07-17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hal000004</td>
<td>2017-07-17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hal000005</td>
<td>2017-07-17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hal000006</td>
<td>2017-07-17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hal000007</td>
<td>2017-07-17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hal000008</td>
<td>2017-07-17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Creation of Metadata Spreadsheets**

**Spreadsheets and More Spreadsheets**
Photo Editing & OCR

- Software like Adobe Photoshop is useful for cropping photos and color correction.
- For books and other multi-page materials, the edited images should be put together into one PDF file on Adobe Acrobat.
- Optical Character Recognition (OCR)-Searchable text in a digital file
  - Can be done via Adobe Acrobat in multiple languages.
Oral History Preservation

- Oral histories being collected during the grant period should be recorded in digital formats like .wav or .mov.
- Oral history narrators should always sign consent forms, and these consent forms should be digitized and archived alongside oral histories themselves.
- Mics should always be used, and sound edited later with software like Premiere Pro and Audacity.
- Some partners have oral histories on legacy formats
- Metadata creation
- Transcription—Can be done in FIU Digital Scholar Studio with Express Scribe Transcription, Adobe Premiere Pro, Final Cut Pro X, or Transcriva.

Oral History Equipment

- Lavalier Microphone
- Camcorder
- Headphones
- Computer
- Audio Recorder
Who do I ask for help?

- You should always consult with your on-site supervisor at your respective institution first.
  - Some of the institutions already have archivists on staff.
- If no one on-site knows the answer or the activity in question is taking place on campus, feel free to contact myself or Enrique. Enrique works with interns to record and edit oral history videos.
  - Training Videos on Oral History Equipment Use:
    - Tascam Audio Recorder: [https://www.youtube.com/watch?v=7jThjDgOs4&t=19s](https://www.youtube.com/watch?v=7jThjDgOs4&t=19s)
    - Other equipment: [https://www.youtube.com/watch?v=2I6_JRfsr9g&t=27s](https://www.youtube.com/watch?v=2I6_JRfsr9g&t=27s)
- There are other personnel on this grant that work with digital collections: Jamie Rogers, Rebeca Bakker, and Kelley Rowan of FIU Digital Collections and Isabel Brador of The Wolfsonian-FIU. We all work collaboratively on the preservation aspects of this grant.

Katie Coldiron, Digital Archivist

- Cell phone: +1 606 923 6887
- Email: kcoldiro@fiu.edu
- Office: MMC, PC 246C
Conclusion: This is a marathon relay, not a sprint.

- This grant lasts for 3 years in total.
  - Most of you will not be working for the entire grant period.
- All of you will be completing integral pieces of the grant.
- You will come out of this experience with skills that will help you on the job market, furthering your education, and beyond.
Works cited

https://guides.lib.utexas.edu/metadata-basics

https://copyright.columbia.edu/basics/fair-use.html


https://www.copyright.gov/what-is-copyright/